Semantic Segmentation Using Label Propagation

Aravindh Mahendran
amahend1@andrew.cmu.edu

Nitish Thatte
nitisht@andrew.cmu.edu

Adwait Gandhe
agandhe@andrew.cmu.edu

Abstract

1 Introduction

Semantic segmentation is the process of assigning a class label to each pixel of the image. This is an important problem in computer vision for understanding the underlying information in an image. While classical segmentation techniques group together the pixels based on low level features, semantic segmentation adopts a supervised learning approach. There are two common approaches for semantic segmentation. The first makes use of low level features and combines them with a learning framework to obtain higher level labels. The second approach is to use low level cues, rather than features, with random fields and learn a unified framework using low level segmentation. In this paper, we first estimate the probability for every pixel label pair by using the k-Nearest Neighbors and the Bag of Words model. A Random Forest [**?**] using the color and texton features is trained which decides whether two connected super pixels share a label and the direction in which the label is propagated.

This paper is structured as follows: Section 1.1 discusses the problem statement. Section 1.2 talks about the related work. Section 2 discusses our method for semantic image segmentation. The experiments conducted and the results obtained are presented in section 3. Finally we summarize our conclusions in section 4.

1.1 Problem Definition

1.2 Related Work

2 Proposed Method

2.1 Attempted Methods

2.2 Intuition

2.3 Our Approach

2.3.1 Features

Textons Textons represent texture information. A collection of filters is used to process each pixel to obtain a per pixel feature vector. This is quantized to a single number by selecting a vector nearest to the current one from a precomputed dictionary. If each vector in the dictionary is considered to be a word, this representation gives one word per image pixel and is called the TextonMap or WordMap. A bag of words histogram is computed for each super pixel or segment by counting word occurrences. We used the filter bank and dictionary provided by [2].
Color Color information is important for separating classes that lack texture or have similar texture. We use the per channel color histogram, color mean and standard deviation from the HSV color space.

2.3.2 Label Propagation

Super pixels and connectivity graph Each image is divided into super pixels at two different scales such that the a coarse super pixels align with corresponding finer ones. This is achieved by using [].A graph is constructed where each fine super pixel is connected to its 4-connected neighborhood and to all fine super pixels inside the same coarse super pixel. We believe that this connectivity helps model non local properties which are important for determining semantically relevant class labels.
Training the classifier We train a classifier that decides whether or not two super pixels should have the same label. We assign a label for each super pixel by taking the majority vote over ground truth labels of each pixel in it. Each edge in the super pixel graph of a training image is considered a positive sample if corresponding super pixels have the same label, negative other wise. This training data is collected over the k nearest neighbors to the test image and used to train a random forest. If done this way, we have too many positive samples and too few negatives and discriminative learning fails to separate them. We resample the training data to ensure a low enough ratio between these before training our classifier. Note that we train a new random forest for every test image. Texton and color features are concatenated to describe each super pixel. Further, each edge in the graph is described by the absolute difference of their features.
Propagating Labels A different segmentation algorithm (see section 2.3.3) gives a probability for each pixel label pair. The entropy for each super pixel can be calculated from this distribution. The label with highest probability is treated as the label assigned to each pixel. A majority vote over this assignment is used to determine the label for each super pixel. When the baseline is inconsistent with the prediction of our random forest, we propagate labels from the super pixel with higher entropy to the super pixel with lower entropy. This propagation is run from the lowest entropy super pixel to the highest entropy super pixel propagating labels to super pixels further down in this sorted list.

2.3.3 Baseline

The label propagation approach discussed above requires per pixel values for $P(Label)$. The algorithm which provides this is referred to as baseline and described below in detail. **Training** Each training image is divided into segments using Meanshift based EDISON [1]. A normalized texton histogram (bag of words representation) is computed for each segment. The ground truth label for each segment is selected by taking a majority vote over per pixel ground truth labels.
Testing For each test image, meanshift segments and normalized texton histograms of the same are computed. The probability distribution $P(Label|feature)$ is computed by counting the occurrences of each label in the K nearest segments from the training data. This distribution is copied to each pixel in the image and a per pixel distribution and entropy is thus available.

3 Experiments

3.1 Description of Experiments

We tested our algorithm on the MSRC V1 Dataset [3].

3.2 Observations

4 Conclusion

References

[1] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):603 –619, may 2002.

[2] Tomasz Malisiewicz and Alexei A. Efros. Recognition by association via learning per-exemplar distances. In *CVPR*, June 2008.

[3] J. Winn, A. Criminisi, and T. Minka. Object categorization by learned universal visual dictionary. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1800 –1807 Vol. 2, oct. 2005.